University of California, Berkeley
Physics 105 Fall 2000 Section 2 (*Strovink*)

SOLUTION TO PROBLEM SET 1

Solutions by T. Bunn and M. Strovink

Reading:
105 Notes 1.1, 1.2, 1.3, 1.4, 1.5.
Hand & Finch 7.1, 7.2, 7.3, 7.4, and 8.7 (pp. 300-302 only).

1. A matrix A is called *orthogonal* if

$$A^{-1} = A^t \ ,$$

where

$$\left(A^t\right)_{ij} \equiv A_{ji} \ .$$

(a) Prove that the product of two orthogonal matrices is also orthogonal.

Solution: Note: We're going to use the summation convention all the time in these solution sets, so you'd better get used to it. Just remember, repeated indices get summed over, unless specifically indicated otherwise. For instance, when we say $(AB)_{ij} = A_{ik}B_{kj}$, the right-hand side means $\sum_{k=1}^{n} A_{ik}B_{kj}$. It's really not so hard.

Also, we've explained the reasoning used in problem 1(a) in more detail than we usually will, for those of you who aren't used to matrix manipulations. Make sure you understand each step of this reasoning.

Suppose A and B are orthogonal, so their transposes are the same as their inverses. Then we need to show that $(AB)^t = (AB)^{-1}$. In other words, we want to show that $(AB)(AB)^t$ is the identity matrix. Well, here goes: The i, j element of this product is

$$\begin{aligned}
((AB)(AB)^t)_{ij} &\overset{\mathbf{1}}{=} (AB)_{ik}(AB)^t_{kj} \\
&\overset{\mathbf{2}}{=} (AB)_{ik}(AB)_{jk} \\
&\overset{\mathbf{3}}{=} A_{il}B_{lk}A_{jm}B_{mk} \\
&\overset{\mathbf{4}}{=} A_{il}A_{jm}B_{lk}B^t_{km} \\
&\overset{\mathbf{5}}{=} A_{il}A_{jm}(BB^t)_{lm} \\
&\overset{\mathbf{6}}{=} A_{il}A_{jm}\delta_{lm} \\
&\overset{\mathbf{7}}{=} A_{il}A_{jl} \\
&\overset{\mathbf{8}}{=} A_{il}A^t_{lj} \\
&\overset{\mathbf{9}}{=} \delta_{ij} \ .
\end{aligned}$$

But δ_{ij} is the i, j element of the identity matrix, so we've shown what we intended to.

Now, as promised, here's an explanation of each "=" sign in the above: (1) This is the ordinary rule for matrix multiplication. (2) The "t" is for "transpose": It means switch the order of the two indices k and j. (3) Matrix multiplication rule applied to each of the two terms in parentheses. (4) Transpose. (5) Matrix multiplication rule (applied in reverse this time). (6) B is orthogonal, so BB^t is the identity, and the l, m element of the identity is δ_{lm}. (7) Substitution rule for the δ symbol. (8) Transpose. (9) A is orthogonal.

(You don't need to write out all of these steps in so much detail. Also, if you know an easier way to do it, that's fine. In particular, if you know that the transpose of AB is B^tA^t, there's a really easy way to do part (a): $(AB)(AB)^t = ABB^tA^t = ABB^{-1}A^{-1} = AA^{-1} = I$.)

(b) Show that if A is a 3×3 orthogonal matrix,

its three column vectors are mutually perpendicular and of unit length.
Solution: Define $\mathbf{V}^j$ to be the j^{th} column vector of A, *i.e.* $(V^j)_i = A_{ij}$. Then

$$\begin{aligned}\mathbf{V}^j \cdot \mathbf{V}^k &= (V^j)_i (V^k)_i \\ &= A_{ij} A_{ik} \\ &= (A^t)_{ki} A_{ij} \\ &= (A^t A)_{kj} \\ &= I_{kj} \\ &= \delta_{kj} \ ,\end{aligned}$$

where I is the identity matrix.

2. Suppose that a vector $\mathbf{x}'$ in the space axes is related to a vector $\mathbf{x}$ in the body axes by

$$\mathbf{x}' = A\mathbf{x} \ ,$$

where A is a transformation matrix. Given a matrix F, find a matrix F', expressed in terms of F and A, such that

$$\mathbf{x}'^t F' \mathbf{x}' = \mathbf{x}^t F \mathbf{x} \ .$$

F and F' are said to be related by a *similarity transformation*.
Solution: If $\mathbf{x}' = A\mathbf{x}$, then $\mathbf{x} = A^{-1}\mathbf{x}'$. (To see why, multiply both sides of the first equation by A^{-1} on the left.) Using the fact that the transpose of AB is $B^t A^t$, we get $\mathbf{x}^t = \mathbf{x}'^t (A^{-1})^t = \mathbf{x}'^t A$. (For the last step, remember that transformation matrices like A are orthogonal.) Now, for any matrix F, $\mathbf{x}^t F \mathbf{x} = \mathbf{x}'^t A F A^{-1} \mathbf{x}'$. (We just substituted for $\mathbf{x}$ and $\mathbf{x}^t$.) So we can choose the matrix F' to be AFA^{-1}.

3. Define the *trace* of a matrix F as

$$\mathrm{Tr}(F) = F_{ij}\delta_{ij} \ ,$$

where, as usual, summation over repeated indices is implied.
(a) Show that $\mathrm{Tr}(F)$ is the sum of the diagonal elements of F.
Solution: Remember the substitution rule for δ_{ij}: If you have an expression containing a δ_{ij}, and the i is being summed over, then you can simply replace it by j and get rid of the δ_{ij}. In this case, the expression $F_{ij}\delta_{ij}$ is the same as F_{jj}. The index j is repeated, so it's being summed over. So F_{jj} means "the sum of all the elements of F that have the same row index and column index." That sounds the same as "sum of all the diagonal elements".

(b) Prove that $\mathrm{Tr}(F)$ is invariant under any similarity transformation.
Solution: We're being asked to show that if F and F' are related by a similarity transformation, then $\mathrm{Tr}(F) = \mathrm{Tr}(F')$. Well, if F and F' are similar, then we can write $F' = AFA^{-1}$ for some A. That means that the i,j element of F' is $F'_{ij} = A_{ik}F_{kl}A^{-1}_{lj}$. So by the trace rule:

$$\begin{aligned}\mathrm{Tr}(F') &= A_{ik}F_{kl}A^{-1}_{lj}\delta_{ij} \\ &= A_{ik}F_{kl}A^{-1}_{li} \\ &= (A^{-1}A)_{lk}F_{kl} \\ &= \delta_{lk}F_{kl} = \mathrm{Tr}(F) \ .\end{aligned}$$

4.
(a) Use the Levi-Civita density ϵ_{ijk} to prove the *bac cab rule*

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b}) \ .$$

Solution:

$$\begin{aligned}(\mathbf{b} \times \mathbf{c})_k &\equiv \epsilon_{klm} b_l c_m \\ (\mathbf{a} \times (\mathbf{b} \times \mathbf{c}))_i &\equiv \epsilon_{ijk} a_j (\mathbf{b} \times \mathbf{c})_k \\ &= \epsilon_{ijk} a_j (\epsilon_{klm} b_l c_m) \\ &= \epsilon_{ijk}\epsilon_{klm} a_j b_l c_m \ .\end{aligned}$$

Any nonvanishing term must have i, j, l, and m all different from k, but every index must take one of only three values. The possibilities are:

- $i = l$, $j = m$: the nonvanishing term $(k \neq i, k \neq j)$ has value $+1$, because $\epsilon_{ijk} = \epsilon_{klm}$ are cyclic permutations of each other.
- $i = m$, $j = l$: the nonvanishing term $(k \neq i, k \neq j)$ has value -1, because $\epsilon_{ijk} = \epsilon_{klm}$ are *not* cyclic permutations of each other.

Changing these words into an equation in the second line below,

$$\begin{aligned}(\mathbf{a}\times(\mathbf{b}\times\mathbf{c}))_i &= \epsilon_{ijk}\epsilon_{klm}a_j b_l c_m \\ &= (\delta_{il}\delta_{jm}-\delta_{im}\delta_{jl})a_j b_l c_m \\ &= a_j b_i c_j - a_j c_i b_j \\ &= b_i(a_j c_j) - c_i(a_j b_j) \\ &= b_i(\mathbf{a}\cdot\mathbf{c}) - c_i(\mathbf{a}\cdot\mathbf{b})\ .\end{aligned}$$

This is bac cab rule for the i^{th} component. Nothing is special about this component, so the rule is proved in general.

(b) Use the bac cab rule to show that

$$\mathbf{a} = \hat{\mathbf{n}}(\mathbf{a}\cdot\hat{\mathbf{n}}) + \hat{\mathbf{n}}\times(\mathbf{a}\times\hat{\mathbf{n}})\ ,$$

where $\hat{\mathbf{n}}$ is any unit vector. What is the geometrical significance of each of the two terms in the expansion?
Solution: by the bac cab rule,

$$\begin{aligned}\hat{\mathbf{n}}\times(\mathbf{a}\times\hat{\mathbf{n}}) &= \mathbf{a}(\hat{\mathbf{n}}\cdot\hat{\mathbf{n}}) - \hat{\mathbf{n}}(\hat{\mathbf{n}}\cdot\mathbf{a}) \\ &= \mathbf{a} - \hat{\mathbf{n}}(\hat{\mathbf{n}}\cdot\mathbf{a}) \\ \mathbf{a} &= \hat{\mathbf{n}}(\mathbf{a}\cdot\hat{\mathbf{n}}) + \hat{\mathbf{n}}\times(\mathbf{a}\times\hat{\mathbf{n}})\ .\end{aligned}$$

This expression splits the vector $\mathbf{a}$ into two parts, the first of which is parallel to $\hat{\mathbf{n}}$, and the second of which is perpendicular to it.

5. Consider three vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$.
(a) Show that

$$(\mathbf{u}\times\mathbf{v})\cdot\mathbf{w} = \epsilon_{ijk}u_i v_j w_k\ ,$$

where, as usual, summation is implied.
Solution: In component language, the dot product rule is $\mathbf{u}\cdot\mathbf{v} = u_i v_i$, and the cross product rule is $(\mathbf{u}\times\mathbf{v})_k = \epsilon_{ijk}u_i v_j$. (The only way to see that this rule for cross products is right is to check it explicitly. Let's see that it works for the z-component: $(\mathbf{u}\times\mathbf{v})_3 = \epsilon_{123}u_1v_2 + \epsilon_{213}u_2v_1 = u_1v_2 - u_2v_1$, which is right.)
So $(\mathbf{u}\times\mathbf{v})\cdot\mathbf{w} = (\mathbf{u}\times\mathbf{v})_k w_k = \epsilon_{ijk}u_i v_j w_k$.

(b) If $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ emanate from a common point, show that $|\epsilon_{ijk}u_i v_j w_k|$ is the volume of the parallelepiped whose edges they determine.
Solution: Using the result of (a),

$$\begin{aligned}|\epsilon_{ijk}u_i v_j w_k| &= |\mathbf{u}\times\mathbf{v})\cdot\mathbf{w}| \\ &= |(\mathbf{u}\times\mathbf{v})||\mathbf{w}||\cos\theta_w| \\ &= |\mathbf{u}||\mathbf{v}||\sin\theta_{uv}|\mathbf{w}||\cos\theta_w|\ ,\end{aligned}$$

where θ_{uv} is the angle between $\mathbf{u}$ and $\mathbf{v}$, and θ_w is the angle between $\mathbf{w}$ and the normal to the plane defined by $\mathbf{u}$ and $\mathbf{v}$. The first three factors yield the area of the parallelogram defined by $\mathbf{u}$ and $\mathbf{v}$, and the last two factors yield the height of the parallelepiped whose base is this parallelogram. The product of the base area and the height is the volume of the parallelepiped.

6. In a complex vector space, a matrix U is called *unitary* if

$$U^{-1} = U^\dagger\ ,$$

where

$$\left(U^\dagger\right)_{ij} \equiv U^*_{ji}\ .$$

Show than an *infinitesimal* unitary transformation T (one that is infinitesimally different from the unit matrix) can be written

$$T \approx I + iH\ ,$$

where I is the unit matrix and H is *Hermitian*, *i.e.*

$$H = H^\dagger\ .$$

Solution:

$$\begin{aligned}T &= I + iH \\ T^\dagger &= I^\dagger + (iH)^\dagger \\ &= I + H^\dagger i^\dagger \\ &= I + H^\dagger(-i) \\ &= I - iH^\dagger\ ,\end{aligned}$$

where in the second line we used the fact that, as for the transpose, $(AB)^\dagger = B^\dagger A^\dagger$. Enforcing the unitarity condition on T,

$$\begin{aligned} T^{-1} &= T^\dagger \\ I &= T^{-1}T \\ &= T^\dagger T \\ &= (I - iH^\dagger)(I + iH) \\ &= I + i(H - H^\dagger) + \mathcal{O}(H^2) \\ &\approx I + i(H - H^\dagger) \\ 0 &\approx i(H - H^\dagger) \\ H^\dagger &\approx H \ . \end{aligned}$$

7. Show that $\mathbf{v}$, $\mathbf{p}$, and $\mathbf{E}$ (velocity, momentum, and electric field) are ordinary ("polar") vectors, while $\boldsymbol{\omega}$, $\mathbf{L}$, and $\mathbf{B}$ (angular velocity, angular momentum, and magnetic field) are pseudo ("axial") vectors.
Solution: Obviously $\mathbf{r} = (x_1, x_2, x_3)$, the position vector, changes sign under parity inversion $\mathcal{P}$, which transforms $(x_1 \to -x_1, x_2 \to -x_2, x_3 \to -x_3)$. Therefore $\mathbf{r}$ is a polar vector. Since $\mathbf{v} \equiv d\mathbf{r}/dt$ and $\mathbf{p} = m\mathbf{v}$, $\mathbf{v}$ and $\mathbf{p}$ are also polar vectors. So is the force $\mathbf{F} = d\mathbf{p}/dt$. The force on a test charge q is $\mathbf{F} = q\mathbf{E}$, so $\mathbf{E}$ must also be a polar vector.

However (see 105 lecture notes eq. (1.7)), $v'_{\text{tang}} = \boldsymbol{\omega} \times \mathbf{r}$ implies that $\boldsymbol{\omega}$ cannot change sign under $\mathcal{P}$, inasmuch as v'_{tang} and $\mathbf{r}$ both do change sign. Therefore $\boldsymbol{\omega}$ is an axial vector. Similarly, $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ is an axial vector because both $\mathbf{r}$ and $\mathbf{p}$ change sign under $\mathcal{P}$. Likewise, the magnetic part of the force $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$ requires $\mathbf{B}$ to be an axial vector because $\mathbf{F}$ and $\mathbf{v}$ both change sign under $\mathcal{P}$.

8. Find the transformation matrix Λ, such that

$$x'_i = \Lambda_{ij} x_j \ ,$$

which describes the following (passive) transformation: relative to the space (primed) axes, the body (unprimed) axes are rotated counterclockwise by an angle ξ about a unit vector $\hat{\mathbf{n}}'$ which has direction cosines n'_1, n'_2, and 0 with respect to the x'_1, x'_2, and x'_3 (space) axes, respectively.
Solution: It is somewhat easier to find Λ^t such that $x_i = \Lambda^t_{ij} x'_j$, as is the case for the Euler rotation (105 lecture notes 1.5). We accomplish the transformation from the primed to the unprimed coordinates in three steps:

(i) Rotate the x'_1 and x'_2 axes about the x'_3 axis so that the new "1" direction (call it x''_1) is along the given unit vector $\hat{\mathbf{n}}'$. Call the rotation matrix which accomplishes this A.

(ii) Rotate the x''_2 and x''_3 axes counterclockwise about the x''_1 (or $\hat{\mathbf{n}}'$) axis by the given angle ξ. Call the matrix which accomplishes this B.

(iii) Reverse the rotation (i). This is necessary so that, for example, if ξ is zero there will have been no net change. Call the matrix which accomplishes this C.

We will then have $\Lambda^t = CBA$, or $\Lambda = A^t B^t C^t$. What are the elements of these rotation matrices? Remembering that we are dealing with *passive* rotations (the axes get rotated, not the vectors), rotation A is

$$A = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} ,$$

where ϕ is the angle whose cosine is n'_1 and whose sine is n'_2. (For getting A right, the only tricky part is to figure out where the minus sign goes. It is straightforward to do this by considering the effect of operating A on $\hat{\mathbf{n}}'$ itself: the result should be, and is, entirely in the "1" direction.) Similarly, rotation matrix B is

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\xi & \sin\xi \\ 0 & -\sin\xi & \cos\xi \end{pmatrix} .$$

C is merely the inverse of A: $C = A^t$. Then $\Lambda = A^t B^t A$. Substituting $\cos\phi = n'_1$, $\sin\phi = n'_2$ and multiplying the matrices,

$$\Lambda = \begin{pmatrix} (n'_1)^2 + (n'_2)^2 \cos\xi & n'_1 n'_2 (1 - \cos\xi) & n'_2 \sin\xi \\ n'_1 n'_2 (1 - \cos\xi) & (n'_2)^2 + (n'_1)^2 \cos\xi & -n'_1 \sin\xi \\ -n'_2 \sin\xi & n'_1 \sin\xi & \cos\xi \end{pmatrix} .$$